

UC Berkeley

UC Berkeley Previously Published Works

Title

Don't throw out the sympatric speciation with the crater lake water: fine-scale investigation of introgression provides equivocal support for causal role of secondary gene flow in one of the clearest examples of sympatric speciation.

Permalink

<https://escholarship.org/uc/item/4dx559sp>

Journal

Evolution letters, 2(5)

ISSN

2056-3744

Authors

Richards, Emilie J
Poelstra, Jelmer W
Martin, Christopher H

Publication Date

2018-10-01

DOI

10.1002/evl3.78

Peer reviewed

Don't throw out the sympatric speciation with the crater lake water: fine-scale investigation of introgression provides equivocal support for causal role of secondary gene flow in one of the clearest examples of sympatric speciation

Emilie J. Richards,^{1,2}  Jelmer W. Poelstra,^{1,3}  and Christopher H. Martin¹ 

¹Biology Department, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

²E-mail: ejr@live.unc.edu

³Biology Department, Duke University, Durham, North Carolina 27710

Received November 12, 2017

Accepted July 18, 2018

Genomic data has revealed complex histories of colonization and repeated gene flow previously unrecognized in some of the most celebrated examples of sympatric speciation and radiation. However, much of the evidence for secondary gene flow into these radiations comes from summary statistics calculated from sparse genomic sampling without knowledge of which specific genomic regions introgressed. This tells us little about how gene flow potentially influenced sympatric diversification. Here, we investigated whole genomes of Barombi Mbo crater lake cichlids for fine-scale patterns of introgression with neighboring riverine cichlid populations. We found evidence of secondary gene flow into the radiation scattered across <0.24% of the genome; however, from our analyses, it is not clear if the functional diversity in these regions contributed to the ecological, sexual, and morphological diversity found in the lake. Unlike similar studies, we found no obvious candidate genes for adaptive introgression and we cannot rule out that secondary gene flow was predominantly neutral with respect to the diversification process. We also found evidence for differential assortment of ancestral polymorphisms found in riverine populations between sympatric sister species, suggesting the presence of an ancestral hybrid swarm. Although the history of gene flow and colonization is more complicated than previously assumed, the lack of compelling evidence for secondary gene flow's role in species diversification suggests that we should not yet rule out one of the most celebrated examples of sympatric speciation in nature without a more thorough investigation of the timing and functional role of each introgressed region.

KEY WORDS: adaptive radiation, gene flow, introgression, population genetics, speciation genomics, sympatric speciation.

Impact Statement

Sympatric speciation, the evolution of reproductive isolation without the aid of geographic barriers, is a fascinating process for its illustration of the power of natural and sexual selection alone to create new species. Despite exhaustive searches, only a handful of case studies provide convincing evidence for sympatric speciation in nature. Even in some of the clearest examples of sympatric speciation, the use of genomic data

has revealed more complicated histories of gene flow from geographically separated populations than once thought. Evidence for these complicated histories is typically collected as a single point estimate of the entire genome, which can tell us that there was secondary gene flow but nothing about how that gene flow may have contributed to the evolution of new species. Here, we exhaustively search whole genomes for signatures of secondary gene flow into the adaptive radiation

of Barombi Mbo crater lake cichlids, one of the clearest case studies of sympatric speciation that has recently come under doubt. We characterized genetic and functional diversity in regions of the genome that have experienced gene flow differently among the species to determine the role of gene flow in the speciation process. Very few regions of the genome appear to have experienced differential gene flow and there was no clear evidence that gene flow after initial diversification in the lake brought in essential adaptive genetic variation for ecological and morphological diversity. We conclude that multiple colonizations of the lake before diversification began (i.e., a hybrid swarm) may have contributed more to the radiation than secondary gene flow after initial diversification. We still cannot rule out the possibility of sympatric speciation in one of the most celebrated examples in nature.

Introduction

Sympatric speciation, the endpoint on the speciation-with-gene-flow continuum, is defined as the evolution of reproductive isolation without the aid of geographic barriers under complete panmixia and constant gene flow between diverging populations (Coyne and Orr 2004; Fitzpatrick et al. 2008; Mallet et al. 2009). Sympatric speciation has fascinated evolutionary biologists since Darwin for its illustration of the power of complex interactions between natural and sexual selection to create new species. Despite intense searches, very few case studies have been able to meet the rigorous criteria for demonstrating sympatric speciation in nature (Coyne and Orr 2004; Bolnick and Fitzpatrick 2007). Even in some of the more convincing examples that do meet these criteria, genomic data have revealed more complex evolutionary histories of multiple colonization and repeated gene flow than previously assumed (Papadopolus et al. 2011; The Heliconius Genome Consortium et al. 2012; Geiger et al. 2013; Alcaide et al. 2014; Igea et al. 2015; Malinsky et al. 2015; Martin et al. 2015a; Kautt et al. 2016; Poelstra et al. 2018).

However, much of the support for complicated histories involving repeated gene flow into sympatric radiations comes from genome-wide tests for gene flow (e.g., Lamichhaney et al. 2015; Martin et al. 2015a; Meier et al. 2017). One prediction of models of speciation with gene flow is that divergence between incipient species should be heterogeneous across the genome (Nosil et al. 2008; Feder et al. 2012; Nosil and Feder 2012). Indeed, high heterogeneity in genomic differentiation has been found across the genomes of many recent or incipient sister species (e.g., Jones et al. 2012; Martin et al. 2013; Poelstra et al. 2014; Soria-Carrasco et al. 2014; Malinsky et al. 2015; McGirr and Martin 2016), although other processes besides differential gene flow across the genome can produce similar heterogeneous pat-

terns (Noor and Bennett 2009; Nachman and Payseur 2012; Cutter and Payseur 2013; Cruickshank and Hahn 2014; Guerrero and Hahn 2017; Ravinet et al. 2017). Only a handful of genes may directly contribute to the speciation process whereas the rest of the genome is porous to gene flow while reproductive isolation is incomplete (Wu 2001; Wu and Ting 2004). Therefore, gene flow detected at the genome-wide level from populations outside the sympatric radiation does not by itself constitute evidence that secondary gene flow was involved in the divergence process among incipient species and shaped the radiation.

The Cameroon crater lake cichlid radiations are some of the most compelling cases for sympatric speciation in the wild (Coyne and Orr 2004; Bolnick and Fitzpatrick 2007; Martin 2012). The most speciose of these radiations is found in the isolated 2.3 km wide volcanic crater lake Barombi Mbo (Trewavas et al. 1972; Schliewen et al. 1994; Schliewen and Klee 2004). Barombi Mbo hosts a radiation of 11 endemic cichlid species, many of which have clear morphological and ecological separation from other sympatric species (Schliewen et al. 1994). Some endemics have evolved unique specializations, such as the spongivore *Pungu maclareni*, seasonal lekking species *Myaka myaka* (C.H.M. pers. obs.), and the deep-water hypoxia specialist *Konia dikume* (Trewavas et al. 1972). Other endemic species, such as *Stomatepia mariae* and *S. pindu*, appear to be incipient or stalled species complexes with only slight morphological and ecological divergence at the extremes of a unimodal distribution of phenotypes (Martin 2012). However, evidence of differential introgression, weak support for Barombi Mbo monophyly, and differences in levels of shared ancestry with outgroup riverine populations from genome-wide RADseq data all suggest secondary gene flow into the radiation after the initial colonization, casting doubt on one of the best examples of sympatric speciation in the wild (Martin et al. 2015a).

Here, we dissect those signals of repeated gene flow to investigate their role in the radiation using whole-genome sequences. We performed exhaustive searches for all genetic patterns consistent with secondary gene flow into the ancestral Barombi Mbo population or into subclades after their initial divergence. We used a machine learning approach to finely dissect and partition different phylogenetic signals across the genome and sliding window genomic scans to test for differential introgression. We found evidence of both shared introgression between sister species and across subclades in the radiation as well as differential introgression among sister species across small regions of the genome. However, functional and genetic diversity in these regions do not paint a clear picture of how introgressed variants may have contributed to speciation in these groups. Our results suggest that either 1) rare introgression of variants in poorly characterized genetic pathways contributed to the morphological and ecological diversity of the radiation (speciation with an allopatric phase),

2) secondary gene flow was predominantly or completely neutral and did not contribute to diversification in Barombi Mbo (sympatric speciation with gene flow), or 3) multiple colonizations of the lake before diversification brought in genetic variation that was then differentially sorted among incipient species (sympatric speciation from a hybrid swarm).

Methods

SAMPLING AND GENOME SEQUENCING

We sequenced whole genomes of one to three individuals from 10 out of the 11 species within the sympatric radiation of Oreochromine cichlids in Cameroon crater lake Barombi Mbo (excluding *Sarotherodon steinbachi*, which is morphologically and ecologically similar to the other three *Sarotherodon* species), an endemic *Sarotherodon* species pair from Lake Ejagham, and outgroup *Sarotherodon* individuals from all three river drainages surrounding the lake: Meme, Mungo, and Cross rivers (Fig. 1A). Details on the collection, extraction, alignment to the *Oreochromis niloticus* reference genome (Brawand et al. 2014), and variant calling protocols following the standard GATK pipeline (McKenna et al. 2010) are provided in the supplementary methods.

CHARACTERIZATION OF INTROGRESSION PATTERNS ACROSS THE GENOME

We exhaustively searched our genomic dataset for patterns of non-monophyletic Barombi Mbo relationships using the machine learning program *Saguaro* (Zamani et al. 2013) to identify regions of the genome that supported phylogenetic topologies consistent with expectations from multiple colonizations and secondary gene flow into the radiation (i.e., paraphyletic/polyphyletic Barombi Mbo radiations). This method infers relationships among individuals in the form of genetic distance matrices and assigns segments across the genomes to different topologies without a *priori* hypotheses about these relationships. We partitioned the genome into a total of 75 unique topologies (well past the inflection point at 30 topologies in which the percentage of the genome covered by each additional topology plateaus; Fig. S1) to exhaustively search for relationships in which subclades or individual Barombi Mbo species were more closely related to riverine populations than other species in the crater lake, suggesting secondary gene flow into this subclade (i. e., introgression) or differential sorting of ancestral polymorphism that originated prior to colonization or after a hybrid swarm. Because we instructed *Saguaro* to propose trees beyond the normal stopping rules based on the percentage of the genome explained by additional topologies to exhaustively search for introgressed regions, the genomic proportions assigned to each tree may not be accurate. These proportions are therefore treated with caution and only used to complement the other results

in this study. Details on the *Saguaro* analysis and filtering strategies for calculating proportions are provided in the supplementary methods.

We also characterized differential introgression within subclades of the radiation on both a genome-wide and local level. We tested for differential introgression between Barombi Mbo species and riverine populations at the genome-wide level using f_4 statistics (Reich et al. 2009; Patterson et al. 2012; Pickrell and Pritchard 2012). We focused on tests of introgression with the three surrounding riverine populations of the most closely related outgroup *Sarotherodon galilaeus* in the nearby Mungo and Meme rivers (MM) and *S. galilaeus* from the more distant Cross River (CR). Based on the tree ((P1, P2), (*S. galilaeus* MM, *S. galilaeus* CR)), f_4 statistics were calculated for combinations of species among (a) *Stomatepia*, (b) the *Konia* + *Pungu* subclade, and (c) *Myaka myaka* with *S. linnelli* as a representative of its sister *Sarotherodon* group. This subset of groupings was chosen to make these analyses more tractable by focusing on species with unique trophic ecologies within the radiation.

Genome-wide f_4 statistics were calculated using the fourpop function in *TreeMix* (Pickrell and Pritchard 2012). Standard error was estimated by jackknifing in windows of 1000 adjacent SNPs to account for linkage disequilibrium. We also visualized the directionality of genome-wide introgression detected with the f_4 statistics using *TreeMix* (v 1.13) (Pickrell and Pritchard 2012). *TreeMix* estimates a maximum likelihood phylogeny of the focal populations and then fits a user-specified number of migration edges to the tree by comparing genetic covariances among populations. We ran *TreeMix* with *S. galilaeus* as root, and with 0 through 20 migration edges. To determine the most likely number of migration events, we performed likelihood-ratio tests comparing each graph to one fewer migration event, starting with 1 versus 0 events, and took the first nonsignificant comparison as the most likely number of migration events.

We investigated whether observed signatures of differential introgression at the genome-wide level contributed variation potentially important to the divergence between species using the f_d statistic, a test statistic similar to f_4 , but designed to be calculated across sliding genomic windows (Martin et al. 2015b). The f_d statistic, a modified version of the *D*-statistic, looks at allele frequencies fitting two allelic patterns referred to as ABBA and BABA based on the tree ((P1,P2),P3,O)), where O is an outgroup species in which no gene flow is thought to occur with the other populations (Martin et al. 2015b). We used individuals of *Coptodon kottae* from another Cameroon crater lake (Fig. 1A) as our distantly related outgroup population for this test and focused on introgression between Barombi Mbo species and surrounding riverine populations *S. galilaeus* MM and *S. galilaeus* CR. Based on the tree ((P1,P2), (*S. galilaeus*, *C. kottae*)), the f_d statistic was calculated for the same combinations of populations used in the

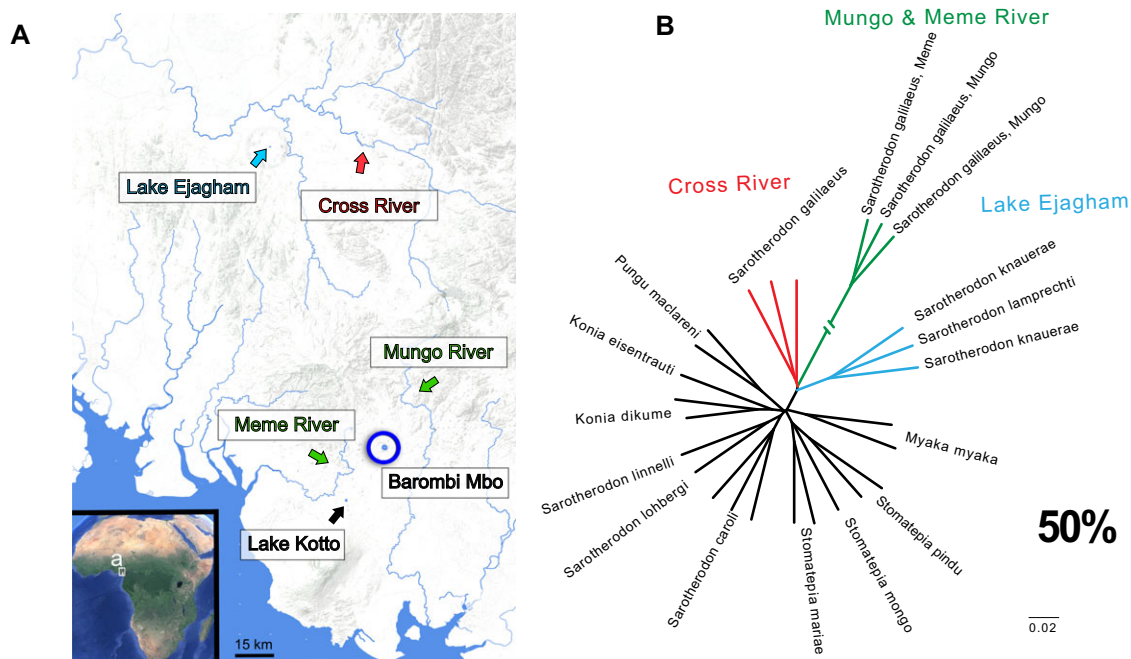


Figure 1. The predominant phylogenetic relationship among Barombi Mbo radiation and neighboring riverine populations. A) Map of lakes and riverine drainages within the volcanic belt of Cameroon in the Northwest and Southwest provinces (Ambazonia). Modified from Figure 2 in Martin et al. (2015a). B) The topology assigned to the largest percentage of the genomes across Barombi Mbo radiation and outgroup *Sarotherodon* species. Across most of the genome Barombi Mbo species (black) are more closely related to each other than riverine outgroup populations of *S. galilaeus* Mungo and Meme River (green) and *S. galilaeus* Cross River (red), or the Lake Ejagham *Sarotherodon* radiation (blue).

genome-wide tests in 50-kb sliding windows with a minimum of 100 variant sites with no missing data within a population using the ABBABABA.py script (Martin et al. 2015b, available on https://github.com/simonhmartin/genomics_general). Significance of f_d values in sliding windows across the genome was evaluated using simulations with no migration using *ms-move* (Garrigan and Geneva 2014). We used the multi-species Markovian coalescent approach (MSMC: (Schiffels and Durbin 2014)) to estimate effective population size changes through time in the Barombi Mbo cichlids and riverine outgroups to parameterize our simulations. Candidate regions were those containing f_d values above the values calculated from simulated windows. For more details on the sliding window calculations of f_d and significance assessment, see supplementary methods.

We divided candidate introgressed regions into three categories: (1) regions found only in a single Barombi Mbo species, (2) regions shared with another species within the same subclade (suggestive of introgression before divergence of the subclade), and (3) regions shared between two or more species across the entire clade, suggesting sympatric speciation after a hybrid swarm (i.e., differential sorting of ancestral polymorphism) or secondary gene flow into multiple subclades (i.e., introgression). We also looked for overlap between introgressed candidate regions be-

tween tests involving *S. galilaeus* MM and *S. galilaeus* CR, another pattern suggestive of a hybrid swarm after initial colonization. For each of these regions, we looked for annotated genes using the well annotated NCBI *Oreochromis* Annotation Release 102 and searched their gene ontology in the phenotype database *Phenoscape* (Mabee et al. 2012; Midford et al. 2013; Manda et al. 2015; Edmunds et al. 2016) and *AmiGO2* (Balsa-Canto et al. 2016) for functions related to the trophic specializations and observed morphological differences among specialist species, such as skeletal system, circulatory system, metabolism, or pigmentation. It is possible that some of the topologies consistent with introgression with outgroups and introgressed regions from f_d tests stem from introgression from unsampled or extinct populations rather than *S. galilaeus* MM or *S. galilaeus* CR directly; however, this should not change the overall conclusion that secondary gene flow events occurred in the history of the radiation or the functional support for the importance of the introgressed regions that we detected.

COMPARISON OF PATTERNS OF INTROGRESSION TO PATTERNS OF GENETIC DIVERGENCE AND DIVERSITY

Reduced levels of genetic polymorphism in a population may indicate a strong selective sweep. We examined introgressed regions

found in only a single Barombi Mbo species for evidence of selection, suggesting that secondary gene flow brought in variation potentially important for speciation. To examine genetic diversity in candidate introgressed regions, we calculated between-population nucleotide divergence (D_{xy}) and within-population nucleotide diversity (π) for pairwise species comparisons among our Barombi Mbo focal specialist species and the riverine outgroups over the same 50-kb windows as the f_d tests (see supplementary methods for more details on these calculations).

Results

WIDESPREAD EVIDENCE OF POLYPHYLY IN BAROMBI MBO RADIATION SCATTERED ACROSS SMALL REGIONS OF THE GENOME

After conservative filtering of segments to remove uninformative regions (Table S1), the Barombi Mbo cichlid radiation formed a monophyletic group across 53% of the genome whereas only 0.6% was assigned to phylogenies indicating a polyphyletic Barombi Mbo using the machine-learning *Saguaro* approach. These polyphyletic relationships are consistent with many processes, including secondary gene flow, incomplete lineage sorting, strong selection (balancing, divergent, background), and ancestral population structure. The most prevalent phylogeny spanned 38.2% of the genome and featured the expected species phylogeny for this group, in which all Barombi Mbo individuals formed a single clade with distant relationships to outgroup riverine *S. galilaeus* populations (Fig. 1B). The second most prevalent topology (spanning 11.8% of the genome) featured identical evolutionary relationships, except for a much shorter branch leading to *S. galilaeus* Mungo and Meme River. Branch lengths produced by *Saguaro* have no direct interpretation as an evolutionary distance (analogous to a neighbor-joining tree), but may be useful for comparison to similar topologies with different branch lengths. Change to (e.g., indicating regions with higher divergence rates; Zamani et al. 2013).

In the 0.6% of the genome supporting polyphyletic Barombi Mbo relationships, we found evidence consistent with multiple colonizations of the lake. Since we were looking for patterns consistent with secondary gene flow or a hybrid swarm for subclades of the radiation, we focused on topologies where single species or entire subclades were more closely related to outgroups than other Barombi Mbo species, which represented only 0.24% of the genome (Fig. 2; Figs. S2–S4). Some topologies featured an entire subclade (e.g., *Stomatepia*) as a monophyletic group more closely related to the riverine populations than other Barombi Mbo species, a relationship consistent with a hybrid swarm scenario before the diversification of the *Stomatepia* subclade (Fig. 2A–C). Other topologies featured individual species more closely related to outgroup riverine pop-

ulations than sister species (Fig. 2D), a relationship consistent with secondary gene flow into that lineage after the initiation of divergence.

The general pattern of polyphyletic relationships was also found in maximum likelihood phylogenetic analyses of a sample of polyphyletic regions initially estimated using *Saguaro*. However, nearly all *Saguaro* topologies in which a single species was more closely related to outgroups differed from the maximum likelihood tree of the same region, which instead featured several Barombi Mbo species as more closely related to the outgroups. The topologies that were consistent across the two analyses and had greater than 85% bootstrap support for the polyphyletic relationship featured entire subclades of Barombi Mbo as more closely related to riverine populations (Fig. 2). For example, in *Stomatepia* we found topologies that grouped multiple species with riverine populations (Fig. 2A). In the *Konia* + *Pungu* subclade, we saw a similar pattern in a topology in which the entire subclade was sister to the riverine outgroup populations (Fig. 2B). In the zooplanktivore *M. myaka*, we did find a topology in which *M. myaka* was sister to the riverine populations (Fig. 2C), but also a topology where *M. myaka*, along with all the Barombi Mbo *Sarotherodon* species, were sister to the riverine outgroup populations (Fig. 2D).

GENOME-WIDE EVIDENCE FOR DIFFERENTIAL INTROGRESSION INTO THE RADIATION

Consistent with evidence of differential introgression from RAD-seq data (Martin et al. 2015a), genome-wide f_d tests provided evidence of genome-wide differential gene flow between some Barombi Mbo sister species and the outgroup riverine species (Table 1). There was significant evidence of introgression in genome-wide tests in tests involving both *S. pinde* in the *Stomatepia* species complex and the hypoxia specialist *K. dikume* in the *Konia* + *Pungu* subclade. Some species pair combinations within these subclades did not show evidence of differential gene flow, suggesting that sympatric speciation potentially occurred for at least some species within the radiation. For example, there was no significant secondary gene flow detected genome-wide in the tests involving sister species *S. mariae* and *S. mongo* or *M. myaka* and *S. linnelli* (Table 1). These results are similar to previous RADseq genome-wide f_d statistics of the radiation, with significant secondary gene flow detected into Barombi Mbo from tests involving *S. pinde* and *S. linnelli* (Martin et al. 2015a). In the previous study, these tests were only significant when more closely related *Sarotherodon* species from Lake Ejagham were used as outgroups rather than more distantly related riverine populations of *S. galilaeus* from Ghana and the Democratic Republic of Congo. It is possible that the introgression detected previously came from an unsampled or extinct riverine population; more likely, there was introgression from the Cameroon *S. galilaeus*

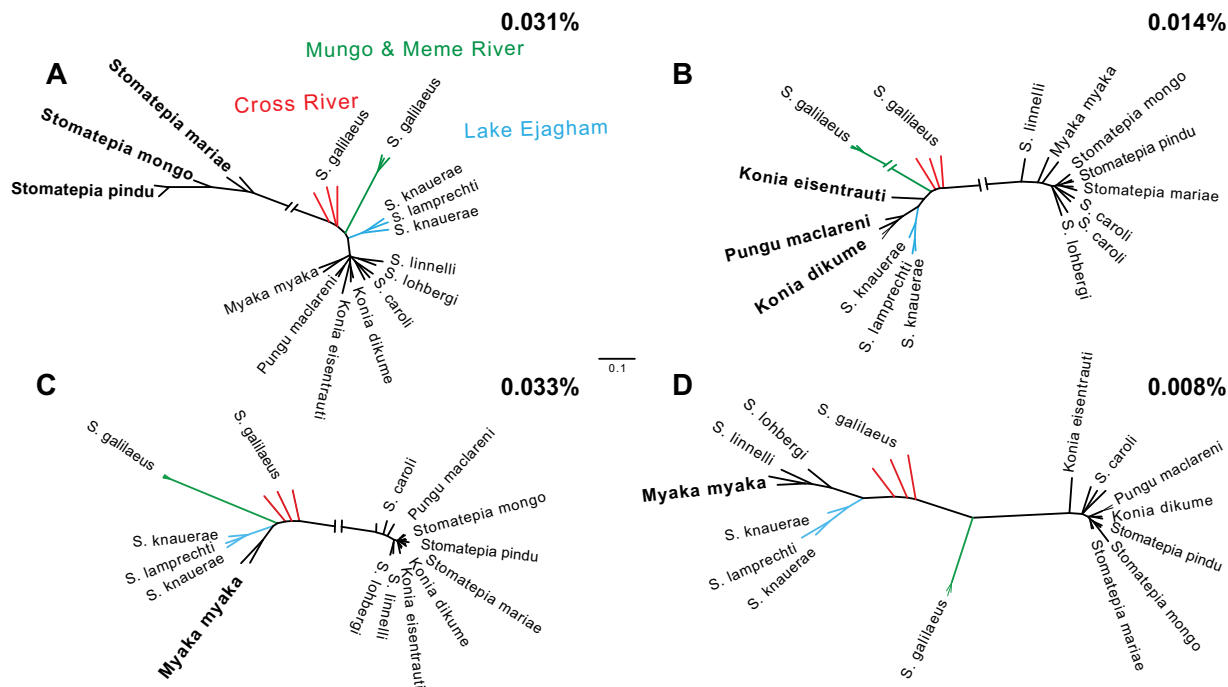


Figure 2. Topologies featuring Barombi Mbo polyphyly with riverine populations involving the Barombi Mbo species with unique ecologies. Across small and independent proportions of the genome (A) the entire *Stomatepia* clade, (B) only *S. pindu*, (C) *M. myaka* and *Sarotherodon* species, and (D) only *M. myaka* were more closely related to outgroups than other Barombi Mbo species. These topologies are consistent with introgression between outgroups and Barombi Mbo and are supported by maximum likelihood analyses. Percentages indicate proportion of the genome assigned to these topologies.

Table 1. Genome-wide f_4 statistics supporting differential introgression within Barombi Mbo radiation. Tests with significant evidence for differential introgression are highlighted in bold. The f_4 statistic was calculated for pairwise combinations among sister species of Barombi Mbo subclades (A, B) and riverine populations of *S. galilaeus* from the Mungo and Meme Rivers (MM) and Cross River (CR).

Introgression with riverine outgroups: (A,B) $\leftarrow \rightarrow$ (MM, CR)			
	f_4 statistic	Z-score	P-value
<i>S. mariae</i> , <i>S. mongo</i>	$-2.04 \times 10^{-7} \pm -5.15 \times 10^{-7}$	-0.39	0.69
<i>S. mariae</i>, <i>S. pindu</i>	$-1.92 \times 10^{-6} \pm -4.48 \times 10^{-7}$	-4.29	1.8×10^{-5}
<i>S. mongo</i>, <i>S. pindu</i>	$-1.59 \times 10^{-6} \pm -4.98 \times 10^{-7}$	-3.19	0.0014
<i>K. dikume</i>, <i>K. eisenrauti</i>	$-2.4 \times 10^{-6} \pm -6 \times 10^{-7}$	-4.01	6.3×10^{-5}
<i>K. eisenrauti</i> , <i>P. maclareni</i>	$-2.12 \times 10^{-7} \pm -6.15 \times 10^{-7}$	0.35	0.73
<i>K. dikume</i>, <i>P. maclareni</i>	$-2.56 \times 10^{-6} \pm -5.86 \times 10^{-7}$	-4.37	1.2×10^{-5}
<i>M. myaka</i> , <i>S. linnelli</i>	$-4.04 \times 10^{-7} \pm -7.11 \times 10^{-7}$	0.56	0.57

riverine populations examined in this study into the Ejagham *Sarotherodon* species examined previously (Martin et al. 2015a).

We also found evidence for widespread gene flow connecting populations across Barombi Mbo and neighboring riverine populations in highly interconnected population graphs using *TreeMix* (Pickrell and Pritchard 2012); the likelihood of each graph did not plateau until reaching 10 admixture events (Fig. 3). On the *TreeMix* population graph with 10 admixture events, gene flow from the Mungo/Meme River populations of *S. galilaeus* occurred directly into individual species *S. mongo* and *K. eisenrauti* rather than the ancestral node of their respective subclades (Fig. S5). The

proportion of admixture inferred for these two events (0.1% into *S. mongo* and 0.4% into *K. eisenrauti*) was similar to the small proportions of the genome assigned to topologies consistent with secondary gene flow in the *Saguaro* analyses. These admixture events pointing to the tips of the graphs suggest secondary gene flow between nearby riverine populations and individual species within the radiation. In all population graphs allowing up to 21 migration events, any admixture from outgroup riverine populations appears to be coming from the Mungo and Meme rivers rather than the Cross River, consistent with the closer geographic proximity of the former drainages.

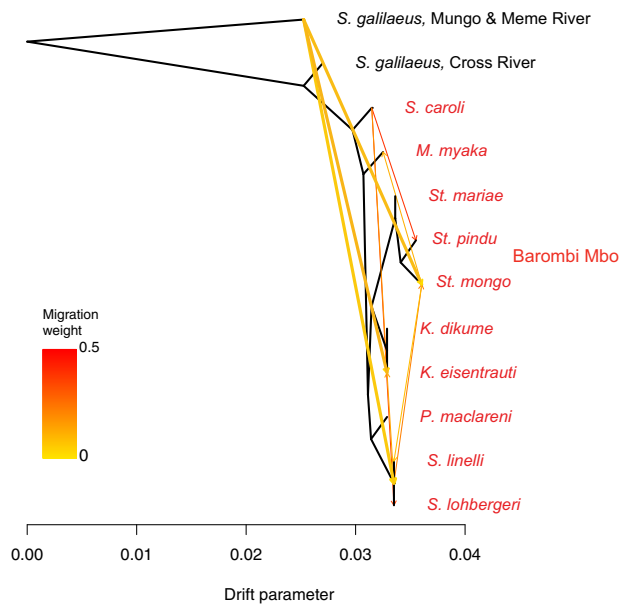


Figure 3. Visualization of genome-wide introgression from riverine *Sarotherodon* populations into Barombi Mbo radiation. *TreeMix* graph illustrating 10 admixture events (with heat colors indicating intensity) on a population graph of the radiation. Admixture events from riverine populations into the radiation are indicated with thicker arrows.

VERY FEW GENOMIC REGIONS CONTAIN SIGNATURES OF DIFFERENTIAL INTROGRESSION INTO INDIVIDUAL BAROMBI MBO SPECIES

We characterized heterogeneity in introgression across the genome by calculating f_d statistics in 50-kb sliding windows among the same combinations of Barombi Mbo species used in genome-wide tests to investigate whether differential introgression contributed variation potentially important in the divergence between species. Very few regions of the genome introgressed into single species from *S. galilaeus* MM (Fig. 4A–C; Fig. S6). Among *Stomatepia* species, only a single candidate introgressed region was found to be differentially introgressed from *S. galilaeus* MM into *S. pindu*, *S. mariae*, and *S. mongo* (50, 80, and 80-kb, respectively), suggesting very little secondary gene flow after initial diversification of *Stomatepia* began (Table 2). Similarly, secondary introgression was also detected for both *Konia* species, *P. maclareni* and *M. myaka* (Table 2). However, only 0.0054–0.1% of the genome introgressed into a single species of a Barombi Mbo subclade from *S. galilaeus* MM, in blocks ranging from 50 to 95-kb in size. Most of the introgressed regions determined from f_d statistics overlapped with regions assigned to introgression topologies from *Saguaro* (Fig. 2), although these topologies made up only a proportion of each of the introgressed regions (0.09–43.7%; Table S2). This incomplete overlap may be driven by the fixed window size of the f_d statistic test, or the

exhaustive genome partitioning in our *Saguaro* approach leading to oversegmentation of the topologies in these regions.

In contrast, a larger proportion of the genome appears introgressed from *S. galilaeus* CR based on our coalescent simulation cut-offs (Fig. S7). Introgression with *S. galilaeus* CR and a single species of Barombi Mbo was detected in 0.7–5% of the genome. These candidate introgressed regions ranged in size from 50 to 220-kb. The size range of introgressed regions was larger in tests with *S. galilaeus* CR than *S. galilaeus* MM (Fig. S8). Larger blocks of introgressed regions could indicate that gene flow from Cross River was more recent than Mungo or Meme River. Absolute genetic divergence between Barombi Mbo species and *S. galilaeus* CR is about seven times smaller across the genome than between Barombi Mbo and *S. galilaeus* MM (Fig. S9–S11; Table S3). It is difficult to assess whether all these candidate introgressed regions from *S. galilaeus* CR represent long tracts of introgressed material given that (1) our moderate sequencing depth prevents phasing and (2) the percentage and length of introgression tracts may change based on the timing of gene flow and divergence in this system. Such knowledge will require demographic modeling of different scenarios of gene flow and its timing, which is still computationally intractable for 11 species. Given this, we focus on the candidate introgressed regions involving *S. galilaeus* MM population. We also use the *S. galilaeus* CR introgressed regions that overlap with the *S. galilaeus* MM introgressed regions to look for patterns of secondary gene flow consistent with a hybrid swarm.

EVIDENCE FOR A HYBRID SWARM

We found shared introgression signals between species both within and among subclades of Barombi Mbo (Fig. 3, Table S4 and S5) in which multiple species shared introgressed regions from a riverine population. This suggests that some of this introgression may have occurred in the ancestral stages of the radiation and differentially sorted among species as they diversified. A majority of these (19/27) were also candidate introgression regions from *S. galilaeus* CR, in which one or more Barombi Mbo species are more closely related to *S. galilaeus* MM while other Barombi Mbo species are more closely related to *S. galilaeus* CR (Table 2). This pattern is suggestive of a hybrid swarm scenario from multiple colonizations by different riverine populations before diversification and the sorting of these ancestral polymorphisms among incipient Barombi Mbo species.

CANDIDATE GENES FOR ADAPTIVE DIVERSIFICATION WITHIN INTROGRESSIONED REGIONS

Although we found evidence of differential introgression among sister species scattered across a small proportion of the genome, the types of genes found in these regions do not paint a clear picture of how this variation may have contributed to speciation

Table 2. Candidate introgressed regions in Barombi Mbo cichlid radiation. These regions feature significant f_d values between riverine populations of *S. galilaeus* (MM: Mungo and Meme River; CR: Cross River) and the three subclades of the radiation focused on in this study. Unannotated regions with no GO terms are marked with (–).

Linkage Group	Position	Gene(s)	Gene Ontology Terms
<i>Myaka myaka</i>			
LG3	17560001–17615000	<i>fcgr2b;fcgr1</i>	IgG binding, immunoglobulin mediated immune response; IgG binding, phagocytosis
LG4	800001–855000	<i>ifi44</i>	defense response to virus; GTP binding
LG5	35530001–35595000	<i>matn4;rbpjl</i>	growth plate cartilage chondrocyte morphogenesis; transcriptional activator activity, RNA polymerase II proximal promoter
LG6	10670001–10740000	<i>prf1;actb</i>	immunological synapse formation, wide porin channel activity; dense body, focal adhesion
LG6	20935001–21025000	<i>jpt2</i>	cytosol; plasma membrane
LG11	28280001–28375000	<i>iqgap3;ttc24;gnrh2;igdcc3;polr3gl;hfe</i>	calmodulin binding, Ras protein signal transduction; biological process; gonadotropin hormone-releasing activity, reproduction; neuromuscular process controlling balance; nuclear chromatin
LG20	1705001–1755000	<i>atp6ap1;taz</i>	DNA-directed RNA polymerase; iron ion transport
LG20	19570001–19630000	<i>sbk1</i>	Rab GTPase binding; positive regulation of bone resorption; O-acyltransferase activity
NT_167475.1	450001–500000	<i>ier5l;h2dmb1;plgrkt</i>	protein serine/threonine kinase activity
NT_167500.1	930001–1000000	<i>cldn15; il20rb</i>	biological process; ribonuclease H2 complex; positive regulation of plasminogen activation
NT_167568.1	10001–60000	<i>zfp235</i>	cell-cell junction assembly, transforming growth factor beta receptor signaling pathway; interleukin-10 receptor activity
NT_167617.1	320001–400000	<i>zg57;nebl</i>	DNA binding, transcription activity
NT_167716.1	220001–270000	–	unknown; cardiac muscle thin filament assembly
NT_167790.1	255001–270000	–	–
NT_167636.1	140001–190000	–	–
<i>Pungu maclareni</i>			
LG3	14525001–14575000	<i>plac8</i>	negative regulation of apoptotic process, chromatin binding
LG17	9260001–9315000	<i>gimap8</i>	GTP binding, regulation of T cell apoptotic process
NT_167557.1	1410001–1495000	<i>pgbd4;UPC</i>	DNA binding transcription factor activity
NT_167671.1	265001–330000	<i>wsc3; UPC(2)</i>	Rho protein signal transduction
NT_167747.1	45001–330000	<i>nirc3;dclk2</i>	negative regulation of I-kappaB kinase/NF-kappaB signaling; peptidyl-serine phosphorylation
NT_168010.1	30001–80000	–	–
<i>Konia dikume</i>			
LG2	18660001–18710000	<i>hmcn1</i>	basement membrane; fin morphogenesis
NT_168013.1	1–70000	<i>siglec12; UPC(2)</i>	cell adhesion, carbohydrate binding;
<i>Konia eisenrauti</i>			
LG6	22460001–22515000	<i>pitpnc1; vwa7</i>	phosphatidylinositol transporter activity; extracellular region, biological process
NT_167586.1	220001–290000	<i>trappc1;psmb6;ephb4*</i>	TRAPP complex; threonine-type endopeptidase activity; ephrin receptor signaling pathway, heart morphogenesis

(Continued)

Table 2. Continued.

Linkage Group	Position	Gene(s)	Gene Ontology Terms
NT_167663.1	535001–630000	<i>s100p</i> ; <i>s100g</i> ; <i>slc29a2</i> ; UPC	RAGE receptor binding; apical plasma membrane, transition ion binding; nucleoside transmembrane transporter activity;-
<i>Stomatepia mongo</i>			
LG20	150001–230000	<i>plod3</i> ; <i>cyb5d</i> ; <i>clec10A</i>	peptidyl-lysine hydroxylation, procollagen-lysine 5-dioxygenase activity; metal ion binding; carbohydrate binding, adaptive immune response
<i>Stomatepia mariae</i>			
LG16-21	31140001–31220000	<i>cxc2</i> ; <i>faim</i> ; <i>parp9</i> ; <i>parp14</i>	neutrophil chemotaxis; apoptotic process; positive regulation of interferon-gamma-mediated signaling pathway; negative regulation of interferon-gamma-mediated signaling pathway
<i>Stomatepia pindu</i>			
NT_167675.1	605001–655000	<i>cmklr1</i> ; <i>c5ar1</i>	G-protein coupled receptor signaling pathway, chemotaxis; complement component C5a receptor activity

*best candidate region for secondary gene flow contributing to diversification; UPC = uncharacterized protein coding gene

(Table 2). For example, differential introgression in *Stomatepia* occurred in regions with genes involved in immune response (Table 2), with no obvious links to the divergent morphological, ecological, or color patterning traits observed among these species (Martin 2012) nor to those traits normally associated with adaptive radiation in cichlid fishes such as body shape, pharyngeal jaw morphology, retinal pigments, or male coloration (Kocher 2004; Barluenga et al. 2006; Wagner et al. 2012; Brawand et al. 2014; Malinsky et al. 2015; Meier et al. 2017). Similarly, in both the *Konia* + *Pungu* and *Myaka* + *Sarotherodon* subclades, introgressed regions were near genes involved in a large range of biological processes not directly associated with adaptive ecological traits in these species, such as *K. dikume*'s hypoxia tolerance, *P. maclareni*'s spongivory, and *M. myaka*'s zooplanktivory.

The clearest candidate for adaptive diversification showing differential introgression into a single species was found between the shallow-water detritivore *K. eisenrauti* and deep-water hypoxia specialist *K. dikume* (Fig. 5). One introgressed region in *K. eisenrauti* contains *ephb4*, a gene involved in blood vessel formation (Herbert et al. 2009; Kawasaki et al. 2014). *K. dikume*'s deep water specialization includes higher blood volume with higher concentrations of hemoglobin (Green et al. 1973). This region overlapped with a region of the genome assigned to a polyphyletic Barombi Mbo topology that featured both *Konia* species plus *P. maclareni* sister to riverine outgroups by *Saguaro* (Fig. 2B, 5). However, this candidate introgressed region is found within a small scaffold that predominantly exhibits f_d values slightly above our simulation cutoff. False positive f_d values can arise from low nucleotide diversity in a region (Martin et al.

2015b), so even this seemingly convincing candidate gene introgression may be driven by lower genetic diversity in the *K. eisenrauti* sequence compared to *K. dikume* rather than true introgression with *S. galilaeus* MM.

From the perspective of gene annotations, there is better support for the contributions of secondary gene flow to the radiation in the introgressed regions that were shared across multiple species of the radiation (Table S4–S5). For example, a region introgressed into *K. eisenrauti* and *P. maclareni* contains *pafah1b3*, a gene involved in platelet activation activity and spermatogenesis in mice (Prescott et al. 2000; Koizumi et al. 2003; Yan et al. 2003), which while not directly associated with *K. dikume*'s traits of higher blood volume with higher concentrations of hemoglobin, is still associated with blood (Table S4). Another region contains an olfactory receptor gene, *or52e8*. The introgressed region containing *or52e8* has signatures of introgression from *S. galilaeus* MM with *K. dikume* and *M. myaka*, as well as from *S. galilaeus* CR with *K. eisenrauti*, *S. mongo*, and *S. mariae* (Table S5). Olfactory signaling is a key component of mate choice and species recognition in cichlids (Plenderleith et al. 2005; Blais et al. 2009). For instance, introduction of olfactory sensory variation from riverine populations may have been important for the divergence of *S. mariae* and *S. pindu*, which appear to exhibit greater sexual than ecological divergence (Martin 2012). Similarly, introgression of olfactory alleles may have played a key role in triggering rapid speciation in Lake Ejagham *Coptodon* cichlids after an 8000 year waiting period (Poelstra et al. 2018).

Finally, two shared introgressed regions contained genes with known skeletal effects: *pfn1* and *mdn1*. The first has been shown

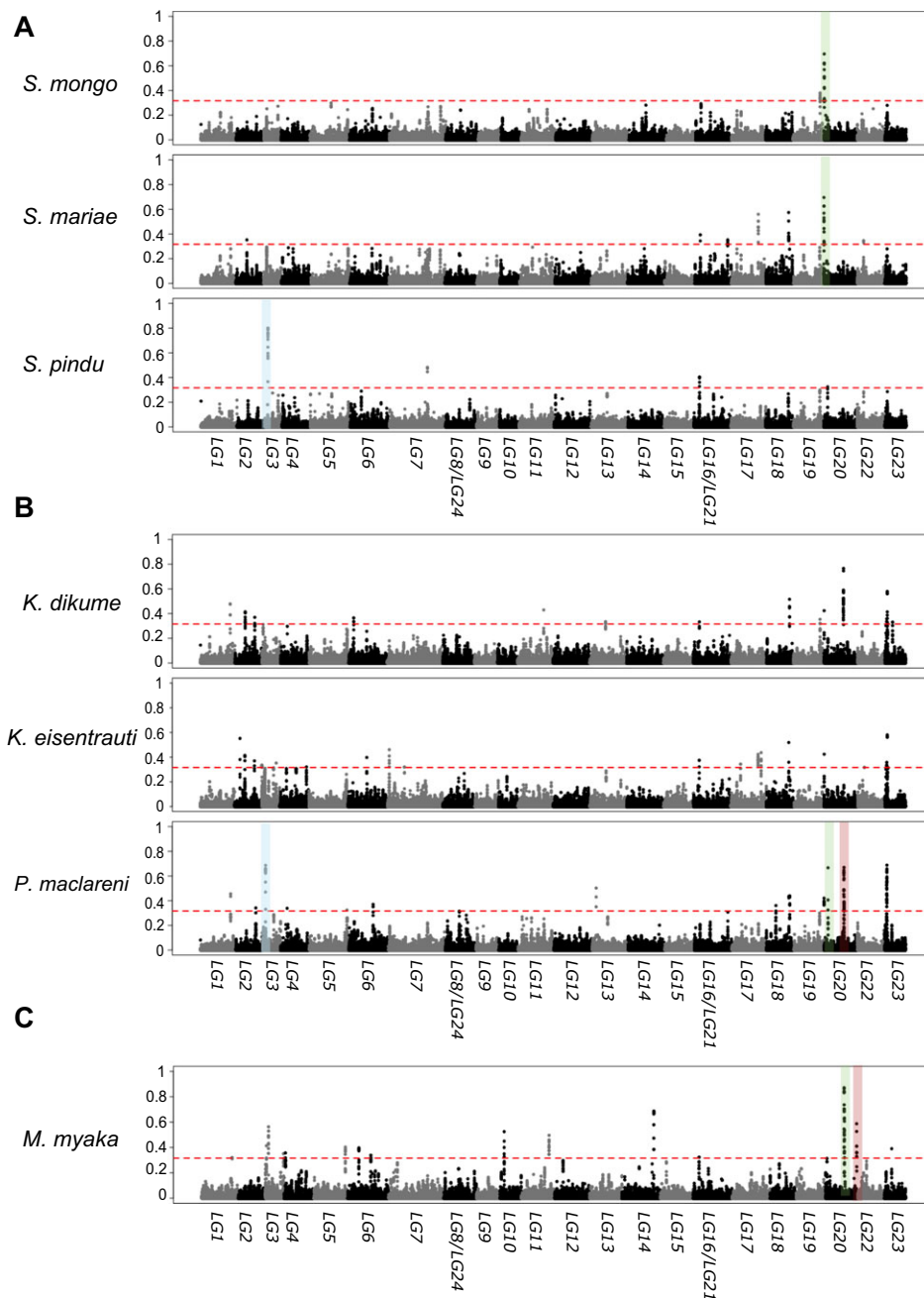


Figure 4. Manhattan plots of f_d values between riverine populations of *S. galilaeus* from Mungo and Meme river and (A) the *Stomatepia* species, (B) combinations of the three species in Konia and Pungu, and (C) *Myaka myaka*. Alternating gray/black colors indicate different linkage groups. Dotted red lines mark the coalescent simulation-based significance thresholds for each test ($f_d = 0.315$). Peaks highlighted in colors represent those signals of introgression shared across different subclades. Manhattan plots for the scaffolds not assigned to the 24 linkage groups are presented in Figure S8.

to effect skeletal development of long bones of limbs in mice (Böttcher et al. 2009) and cell migration during gastrulation in zebrafish (Lai et al. 2008). Mutations in *mdn1* have known craniofacial effects in zebrafish, causing changes in cranium, mandibular arch, and eye size (Kettleborough et al. 2013). The region containing *mdn1* has signatures of introgression from *S. galilaeus* MM with *P. maclareni* and *S. pindu*, as well as introgression from

S. galilaeus CR with *K. dikume* and *S. mongo*. Craniofacial traits such as jaw length, head depth, and orbit diameter are divergent among these Cameroon cichlid species (Martin 2012). For instance, *P. maclareni* has smaller and thicker oral jaws, variation in tooth size, and thicker lips compared to other Barombi Mbo species (Trewavas et al. 1972). *Stomatepia mongo* has a highly elongated neurocranium compared to sister species *S. mariae* and

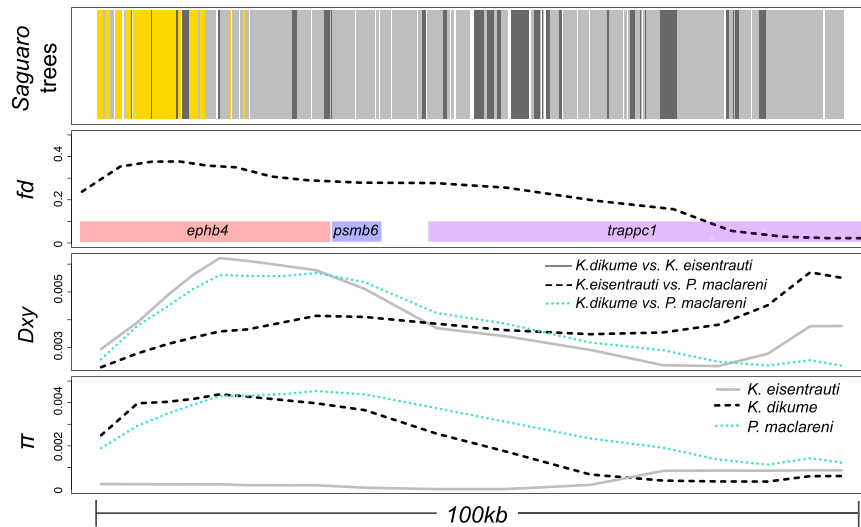


Figure 5. Candidate adaptive introgression region in the *Konia* species pair containing gene *ephb4*. Row 1 shows the topologies assigned by *Saguaro* to the region. Gold blocks were assigned to a topology featuring the *Konia* species pair plus *Pungu maclareni* sister to the riverine *S. galilaeus* populations (Fig. 2B), light grey blocks were assigned to the predominant monophyletic topology (Fig. 1B), and dark grey blocks were assigned to any other topology. Row 2 shows the peak signal of introgression across scaffold NT.167586.1 detected from the f_d statistic across the three-test combination involving *K. dikume*, *K. eisenrauti*, *Pungu maclareni* and riverine populations of *S. galilaeus* MM in overlapping 50-kb windows and the genes in this peak (*ephb4*, *psmb6*). Row 3 shows between-population divergence (D_{xy}) among the combinations of *Konia* and *Pungu* species calculated in overlapping 50-kb windows. Row 4 shows within-population diversity (π) in the same non-overlapping 50-kb windows for the species of *Konia* and *Pungu*. The data from rows highly 2–4 were smoothed using the function `smooth.spline` in R with a `spar` of 0.1 for ease of visualization in the figure.

pindu (Trewavas et al. 1972; Musilová et al. 2014). Signatures of introgression from different riverine populations in regions containing putative adaptive alleles suggest that secondary gene flow from multiple riverine populations prior to diversification (i.e., a hybrid swarm scenario) may have contributed variation important for the radiation.

WEAK SUPPORT FOR SELECTION ON CANDIDATE INTROGRESSED REGIONS

Within the limitations of our sample sizes for each population in this study, the amount of genetic diversity in introgressed regions does not suggest strong divergent selection on introgressed genetic variation due to selective sweeps. In line with the presence of peaks in f_d values in these regions, between-population diversity (D_{xy}) was typically high between one of the species and its sister species (Fig. 4; Table S6). However, within-population diversity across many of these regions was often greater or comparable to scaffold and genome-wide averages (Table S6), suggesting these regions may not have experienced hard selective sweeps that would support their role in adaptive divergence among species. Only a few differentially introgressed regions in *S. pindu* ($n = 1$), *S. mongo* ($n = 1$), and *P. maclareni* ($n = 1$) exhibited genetic diversity an order of magnitude lower than the linkage group average (Table S6), consistent with a selective sweep.

In summary, although we found evidence for differential secondary gene flow between sister species in the radiation, we did not find clear evidence from introgressed genes with well-known roles in adaptive divergence between cichlids nor strong evidence of selection on these regions. However, these alleles may certainly still serve adaptive functions or may be involved in genetic incompatibilities between species. Furthermore, given our sample size limitations, we cannot yet provide a good estimate of selection strength on these regions.

Discussion

EQUIVOCAL EVIDENCE THAT SECONDARY GENE FLOW PROMOTED THE DIVERSIFICATION OF BAROMBI MBO CICHLIDS

Our fine-scale investigations of introgression across the genomes of a celebrated putative example of sympatric speciation are consistent with two possible scenarios: (1) sympatric speciation in the presence of continuous neutral secondary gene flow into the radiation, or (2) speciation initiated by secondary gene flow; our data are not consistent with a scenario involving an extensive period of allopatry and secondary contact promoting speciation. We found little support for the latter allopatric scenario using both *Saguaro* machine learning and sliding-window f_d statistics to exhaustively search for differential introgression into single

species. From the *Saguaro* analyses, our most conservative estimate of introgression into a single species within the radiation ranged from 0.013–0.019% of the genome. Estimates are similarly small for the f_d statistics, ranging from 0.0054–0.1% of the genome from *S. galilaeus* populations in the neighboring Mungo and Meme Rivers. Furthermore, even these significant outliers may represent false positives, particularly in the case of introgression with *S. galilaeus* from Cross River, where 0.7–5% of the genome appears to have introgressed from this population. More complex coalescent simulations, larger sample sizes, and knowledge of the timing of diversification and gene flow events are needed to accurately assess the signatures observed. It is also difficult to distinguish signatures of differential introgression from the biased assortment of ancestral polymorphism into modern lineages; (i.e., a hybrid swarm scenario that is still consistent with sympatric divergence entirely within the crater lake). Finally, even if our statistical outliers represent differentially introgressed regions, their importance to the speciation process is equivocal. Within the limitations of our sample sizes, we found no evidence of selective sweeps in most of these regions that would suggest they aided in divergence between species and most regions contained housekeeping genes that do not clearly suggest how introgressed variation would have contributed to the radiation. Nonetheless, these genes may still serve an adaptive function or could be involved in genetic incompatibilities between diverging sympatric populations, possibly introduced by an initial hybrid swarm in the lake (Seehausen 2004, 2013; Schumer et al. 2015).

EVIDENCE FOR A HYBRID SWARM FURTHER COMPLICATES THE ROLE OF GENE FLOW IN THE SPECIATION PROCESS

Beyond speciation scenarios involving secondary gene flow aiding the completion of speciation, our findings are suggestive of another scenario for sympatric speciation in this system: sympatric speciation from an initial hybrid swarm resulting from the differential sorting of ancestral polymorphisms among incipient species (e.g., see Fig. 1 in Martin et al. 2015a). A hybrid swarm is not easily detectable using the f_d statistic because introgressed variation could be shared among diverging sister species, producing an f_d value near zero due to a BBBA rather than ABBA pattern (Reich et al. 2009; Patterson et al. 2012). However, some of the f_d peaks appear to be (1) shared across at least two of the sister species in a subclade, (2) shared between species of different subclades, or (3) contain variation from both riverine populations (Mungo/Meme and Cross Rivers) that appears differentially sorted among sister species. All three of these patterns are consistent with an ancestral hybrid swarm before divergence between sister species occurred. This pattern of differential sorting of variation from a hybrid swarm from f_d analyses could also

result from a lack of power in this statistic to distinguish the directionality of the introgression detected in those regions when using biallelic patterns and four populations (e.g., when two populations share similar allele patterns, the other two populations can share the opposite allele pattern by default). However, we also found evidence that entire subclades (e.g., *Stomatepia*) were more closely related to riverine populations than other Barombi Mbo subclades using our *Saguaro* analyses, a pattern that is also consistent with a hybrid swarm (e.g., Fig. 2).

It is in these regions that we find the best candidate genes for secondary gene flow contributing to diversification in this system. These candidates are known to affect important traits in other adaptive radiations of cichlids, including olfactory signaling (Nikaido et al. 2013, 2014; Azzouzi et al. 2014; Keller-Costa et al. 2015) and pharyngeal jaw morphology (Muschick et al. 2011; Brawand et al. 2014; Malinsky et al. 2015). These findings are similar to studies on other systems using similar approaches which found compelling cases for adaptive introgression contributing to diversification (e.g., Abi-Rached et al. 2011; The Heliconius Genome Consortium et al. 2012; Huerta-Sánchez et al. 2014; Lamichhaney et al. 2015; Stankowski and Streisfeld 2015; Arnold et al. 2016; Meier et al. 2017), including our own previous work (Richards and Martin 2017). For example, several studies have found convincing candidate genes/variants in introgressed regions suggesting that adaptive introgression played a role in shaping ecological and morphological diversity. These include the detection of introgressed alleles linked to wing-color patterning involved in mimicry and mate selection in *Heliconius* butterflies (The Heliconius Genome Consortium et al. 2012), flower coloration involved in pollinator preferences for *Mimulus* species (Stankowski and Streisfeld 2015), and oral jaw size variation involved in scale-eating trophic specialization in *Cyprinodon* pupfishes (Richards and Martin 2017).

THE CHALLENGES OF SUPPORTING OR REJECTING A ROLE FOR SECONDARY GENE FLOW FROM TESTS OF DIFFERENTIAL INTROGRESSION

There are some caveats to our interpretations of secondary gene flow and its functional role in the ecological and morphological diversity observed within the lake. Recombination rate varies across genomes and determines the scale over which patterns of admixture and differentiation vary (Smukowski and Noor 2011). In our fixed sliding window size of 50-kb, we may have missed important patterns of introgression in regions of recombination hotspots, where such patterns are expected to be very localized (Schumer et al. 2018). Shared variation among species may reflect unsorted polymorphisms from structured ancestral populations rather than hybridization. Introgression events can also be hard to distinguish from ongoing balancing selection of ancestral

polymorphism that is sieved between species (Guerrero and Hahn 2017).

Furthermore, although we focused on searching for genetic signatures of hard selective sweeps, introgressed regions with intermediate to high nucleotide diversity may have undergone soft selective sweeps, in which selection drives multiple adaptive haplotypes to fixation. Some of these introgressed regions may have been adaptive and undergone soft selective sweeps, although the relative contributions of hard sweeps versus soft sweeps during adaptation and speciation is still the subject of much debate (Hermisson and Pennings 2005, 2017; Pritchard et al. 2010; Jensen 2014; Schrider et al. 2015).

Although a search for candidate genes with plausible roles in speciation is a good starting point when considering whether introgressed regions contributed to divergence, there are some caveats to ruling out the importance of secondary gene flow based solely on gene annotations. Not all the effects of a gene may be known, even for model organisms, and there is growing support for an omnigenic model of the link between genotype and phenotype, where all genes expressed in relevant cells can potentially influence a trait (Boyle et al. 2017). Likewise, not all traits involved in species divergence are known and more cryptic traits (e.g., metabolism, physiology) could be as important as the more obviously divergent morphological traits (e.g. McGirr and Martin 2018). Understanding which traits and genes are involved in speciation is now becoming the difficult problem of functional genetic analyses for anything sweeping within a population (see also Richards et al. 2018).

Even if there are no coding regions present in the introgressed region, it may still contain important regulatory elements that affect genes in other regions that underlie speciation traits. It could also house variants that cause genetic incompatibilities and provide a source of postzygotic reproductive isolation. Divergence caused by genetic incompatibilities may not leave signatures of hard selective sweeps in the genome either. One prediction of the hybrid swarm hypothesis is that some of the novel allele combinations introduced by hybridization could include segregating genetic incompatibilities (Seehausen 2004, 2013; Abbott et al. 2013). The presence of DMIs has been shown in other African cichlids (Stelkens et al. 2010), but little is known about what specific regions of the genome contribute to these patterns. Detecting candidate regions of DMIs in non-model organisms is a currently growing body of work mainly investigated in relatively recent hybrid zones (Schumer et al. 2014, 2015, 2018), but is a promising avenue of future research, particularly given the support for a potential hybrid swarm in the evolutionary history of Barombi Mbo shown here.

Although the limitations of this study make it hard to definitively support or reject a role for gene flow in the adaptive radiation of Barombi Mbo, we highlight some of the gaps in knowledge that

need to be filled before we can understand what role the observed secondary gene flow played in diversification. A more thorough understanding of the evolutionary history of the group (e.g., the timing and duration of gene flow events, divergence times among species) will be useful in determining the amount of introgression we would expect to see in the contemporary genomes of the species under different speciation scenarios. A thorough investigation of all the regions of the genome that have undergone soft and hard selective sweeps and the timing of those sweeps, alongside a search for genetic incompatibilities among species, will make it easier to understand the functional importance of observed introgressed regions to diversification. The demographic analyses needed for a radiation of 11 species are still largely intractable, but will eventually provide much-needed insights into the role of the secondary gene flow observed in this classic putative example of sympatric speciation. However, some recently proposed statistics that infer the direction and relative timing of gene flow to speciation provide a promising future avenue for distinguishing between different scenarios of introgression (e.g., hybrid speciation vs. introgression after initial divergence; Hibbins and Hahn 2018) without extensive demographic modeling.

BEST AND WORST REMAINING CASES FOR SYMPATRIC SPECIATION WITHIN THE BAROMBI MBO CICHLID RADIATION

Although the radiation as a whole may not have entirely arisen in sympatry, some sister species within Barombi Mbo may be better case studies of the process than others. Within the three-species *Stomatepia* subclade, there is little evidence that secondary gene flow played an important role in speciation. On a genome-wide level, we detected secondary gene flow for *S. pindu* and *S. mariae*, but not *S. mongo* (Table 1). This is intriguing given that a previous study found continuous unimodal morphological variation connecting the *pindu/mariae* species complex (Martin 2012), but did not include the rare and morphologically distinctive *S. mongo* (Musilová et al. 2014). Furthermore, while the *pindu/mariae* species complex appears unimodal across ecological phenotypes and trophic axes, it exhibits bimodality along the axis of color pattern, indicating that disruptive sexual selection on patterning may be the primary driver and initial barrier to reproduction for this group (Martin 2012). This provides a category of functional traits to search for when trying to determine if secondary gene flow played a role in their current level of divergence.

No reproductive isolating barriers or genetic incompatibilities have yet been quantified among any Barombi Mbo species (in contrast to the strong assortative mating by size, color, morphology, and diet documented in Lake Ejagham *Coptodon* species (Martin 2013). *S. pindu* and *S. mariae* can produce viable F1 hybrids in a no-choice laboratory environment (C.H.M. pers. obs.); however, all Barombi Mbo species appear to avoid

hybridizing when housed in mixed-species laboratory aquaria and no hybrid courtship has been observed in the field (C.H.M. pers. obs.). Estimates from scale growth rings suggest only moderate disruptive selection on ecological traits in the *Stomatepia pindu/mariae* species pair. These species were also not significantly different in their dietary source of carbon or relative trophic position from stable isotope analyses (Martin 2012) and their diets are only marginally distinct based on stomach content analysis, containing predominantly insect/shrimp prey in *pindu* and insect/fish prey in *mariae* (Trewavas 1972; C.H.M. unpublished data). Very little is known about reproductive isolating barriers in other Barombi Mbo cichlids, aside from the interesting and unique temporal isolation of *Myaka myaka* due to its seasonal lekking behavior (C.H.M. pers. obs.).

On a fine scale, the few introgressed regions unique to each of the *Stomatepia* species contained only immune response genes. Shared signatures of introgression among two of the three species or with other Barombi Mbo species represented a larger proportion of the genome than differentially introgressed regions within each species, although both types of introgression were rare across the genome. In those regions that introgressed in more than one *Stomatepia* species, but from different riverine populations, we found genes with functions more typically associated with cichlid radiations. Even for the two monotypic specialist species *M. myaka* and *P. maclareni*, introgressed regions found solely in these species contained mainly housekeeping genes, suggesting secondary gene flow may have been neutral in the evolution of their trophic specializations. However, they also share signatures of introgression in the same regions as *Stomatepia*, suggesting that these monotypic specialists may have obtained variation for their specialized traits from the sorting of ancestral polymorphisms within a hybrid swarm before the radiation began.

Among all the ecologically divergent species pairs focused on in this study, *K. eisenrauti* and *K. dikume* are the least convincing as a putative example of sympatric speciation between sister species. Differentially introgressed regions between *K. dikume* and *K. eisenrauti* include a region containing *ephb4*, involved in heart and blood vessel development. Given *K. dikume*'s hypoxia specialization, this region is the best potential candidate in this study for secondary gene flow and an allopatric phase of speciation contributing to diversification between species in the radiation. These two species also exist in microallopatry; *K. eisenrauti* is an abundant detritivore found only along the shallow littoral region of the lake whereas *K. dikume* is a deep-water specialist on *Chaoborus* midge larvae, which have only been collected in deep-water gill nets (Trewavas et al. 1972; Schlieuwen et al. 1994; C.H.M. pers. obs.). Both species are mouthbrooders and likely breed in non-overlapping habitats although nothing is known about the breeding habits of *K. dikume*.

Conclusion

The complex history of colonization in the Barombi Mbo crater lake cichlid radiation found in this and a previous genome-wide study suggests that secondary gene flow may have played a role in the speciation process, which violates one of the strict criteria for demonstrating sympatric speciation in the wild (Coyne and Orr 2004). Our fine-scale dissection of introgressed regions across the entire genome suggests that supporting or rejecting a role for secondary gene flow in speciation will require an understanding of the functional alleles within each region and their evolutionary history. Nonetheless, we can rule out a scenario in which extensive secondary gene flow after a long allopatric phase, such as reinforcement, contributed to diversification in any Barombi Mbo species. Instead, small and scattered regions of secondary gene flow that were differentially sorted among these sympatric species may have provided variation with undiscovered functional effects on the divergent ecologies and morphologies seen in the lake or this gene flow was predominantly neutral with respect to its role in the speciation process. We found more convincing evidence that secondary gene flow contributed adaptive variation during an initial hybrid swarm within the lake and later sorting of that variation among species in sympatry. Disentangling the effects of a putative hybrid swarm from secondary contact on the speciation process will require a better understanding of the timing of gene flow events compared to the diversification times of Barombi Mbo species. We found evidence for gene flow into the radiation both before and after initial diversification of subclades within the lake. Even without this information, equivocal support for a functional role of secondary gene flow in the radiation of Barombi Mbo cichlids suggests that we should not rule out the possibility of sympatric speciation in this system just yet.

AUTHOR CONTRIBUTIONS

E.J.R. wrote the manuscript and conducted all analyses except for *TreeMix* and MSMC; J.W.P. aligned sequences to the reference genome, called variants, and conducted *TreeMix* and MSMC analyses. C.H.M. collected samples, provided the genomic data, and acquired funding for the study. All authors contributed to the development of ideas presented in the study and revised the manuscript.

ACKNOWLEDGMENTS

This study was funded by a National Geographic Society Young Explorer's Grant, a Lewis and Clark Field Research grant from the American Philosophical Society, NSF CAREER Award #1749764, and the University of North Carolina at Chapel Hill to C.H.M. We gratefully acknowledge the Cameroonian government and the local chiefs of Barombi Mbo village and surrounding communities for permission to conduct this research. We thank Cyrille Dening Touokong, Jackson Waite-Himmelwright, and Patrick Enyang for field assistance; Nono Gonwuou for help obtaining permits; and Katelyn Gould and Logan Buie for help with DNA extractions and quantification.

DATA ARCHIVING

All datasets used for this study are deposited in Dryad (<https://doi.org/10.5061/dryad.712kj7v>) and the NCBI Short Read Archive (Bioproject # PRJNA482609).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

LITERATURE CITED

- Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, et al. 2013. Hybridization and speciation. *J. Evol. Biol.* 26:229–246.
- Abi-Rached, L., M. J. Jobin, S. Kulkarni, A. McWhinnie, K. Dalva, L. Gragert, F. Babrzadeh, B. Gharizadeh, M. Luo, F. A. Plummer, et al. 2011. The Shaping of Modern Human Immune Systems by Multiregional Admixture with Archaic Humans. *Science* 334:89–94.
- Alcaide, M., E. S. C. Scordato, T. D. Price, and D. E. Irwin. 2014. Genomic divergence in a ring species complex. *Nature* 511:83–85.
- Arnold, B. J., B. Lahner, J. M. DaCosta, C. M. Weisman, J. D. Hollister, D. E. Salt, K. Bomblied, and L. Yant. 2016. Borrowed alleles and convergence in serpentine adaptation. *Proc. Natl. Acad. Sci. U. S. A.* 113:8320–8325.
- Azzouzi, N., F. Barloy-Hubler, and F. Galibert. 2014. Inventory of the cichlid olfactory receptor gene repertoires: identification of olfactory genes with more than one coding exon. *BMC Genomics* 15:586.
- Balsa-Canto, E., D. Henriques, A. Gabor, and J. R. Banga. 2016. AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology. *Bioinformatics* 32:1–2.
- Barluenga, M., K. N. Stölting, W. Salzburger, M. Muschick, and A. Meyer. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439:719–723.
- Blais, J., M. Plenderleith, C. Rico, M. I. Taylor, O. Seehausen, C. van Oosterhout, and G. F. Turner. 2009. Assortative mating among Lake Malawi cichlid fish populations is not simply predictable from male nuptial colour. *BMC Evol. Biol.* 9:53.
- Bolnick, D. I., and B. M. Fitzpatrick. 2007. Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* 38:459–487.
- Böttcher, R. T., S. Wiesner, A. Braun, R. Wimmer, A. Berna, N. Elad, O. Medalia, A. Pfeifer, A. Aszódi, M. Costell, and R. Fässler. 2009. Profilin 1 is required for abscission during late cytokinesis of chondrocytes. *The EMBO Journal* 28:1157–1169.
- Boyle, E. A., Y. I. Li, and J. K. Pritchard. 2017. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* 169:1177–1186.
- Brawand, D., C. E. Wagner, Y. I. Li, M. Malinsky, I. Keller, S. Fan, O. Simakov, A. Y. Ng, Z. W. Lim, E. Bezault, et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513:375–381.
- Coyne, J. A., and A. Orr. 2004. *Speciation*. Oxford Univ. Press, Sunderland, MA.
- Cruickshank, T. E., and M. W. Hahn. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* 23:3133–3157.
- Cutter, A. D., and B. A. Payseur. 2013. Genomic signatures of selection at linked sites: unifying the disparity among species. *Nat. Rev. Genet.* 14:262–274.
- Edmunds, R. C., B. Su, J. P. Balhoff, B. F. Eames, W. M. Dahdul, H. Lapp, J. G. Lundberg, T. J. Vision, R. A. Dunham, P. M. Mabey, et al. 2016. Phenoscope: Identifying candidate genes for evolutionary phenotypes. *Mol. Biol. Evol.* 33:13–24.
- Feder, J. L., S. P. Egan, and P. Nosil. 2012. The genomics of speciation-with-gene-flow. *Trends Genet.* 28:342–350.
- Fitzpatrick, B. M., J. A. Fordyce, and S. Gavrillets. 2008. What, if anything, is sympatric speciation? *J. Evol. Biol.* 21:1452–9.
- Garrigan, D., and A. Geneva. 2014. msmove: A modified version of Hudson's coalescent simulator ms allowing for finer control and tracking of migrant genealogies. <https://doi.org/10.6084/m9.figshare.1060474>.
- Geiger, M. F., J. K. McCrary, and U. K. Schliewen. 2013. Crater Lake Apoyo revisited: population genetics of an emerging species flock. *PLoS One* 8:1–17.
- Green, J., S. A. Corbet, and E. Betney. 1973. Ecological studies on crater lakes in West Cameroon The blood of endemic cichlids in Barombi Mbo in relation to stratification and their feeding habits. *J. Zool.* 170:299–308.
- Guerrero, R. F., and M. W. Hahn. 2017. Speciation as a sieve for ancestral polymorphism. *Mol. Ecol.* 26:5362–5368.
- Herbert, S. P., J. Huisken, T. N. Kim, M. E. Feldman, B. T. Houseman, R. A. Wang, K. M. Shokat, and D. Y. R. Stainier. 2009. Arterial-Venous Segregation by Selective Cell Sprouting: An Alternative Mode of Blood Vessel Formation. *Science* 326:294–298.
- Hermisson, J., and P. S. Pennings. 2005. Soft Sweeps. *Genetics* 169:2335–2352.
- Hermisson, J., and P. S. Pennings. 2017. Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods Ecol. Evol.* 8:700–716.
- Hibbins M.S. and M.W. Hahn. 2018. Population genetic tests for the direction and relative timing of introgression. *bioRxiv*. <https://doi.org/10.1101/328575>.
- Huerta-Sánchez, E., X. Jin, Asan, Z. Bianba, B. M. Peter, N. Vinckenbosch, Y. Liang, X. Yi, M. He, M. Somel, et al. 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512:194–197.
- Igea, J., D. Bogarín, A. S. T. Papadopoulos, and V. Savolainen. 2015. A comparative analysis of island floras challenges taxonomy-based biogeographical models of speciation. *Evolution* 69:482–491.
- Jensen, J. D. 2014. On the unfounded enthusiasm for soft selective sweeps. *Nat. Commun.* 5:5281.
- Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61.
- Kautt, A. F., G. Machado-Schiaffino, J. Torres-Dowdall, and A. Meyer. 2016. Incipient sympatric speciation in Midas cichlid fish from the youngest and one of the smallest crater lakes in Nicaragua due to differential use of the benthic and limnetic habitats? *Ecol. Evol.* 6:5342–5357.
- Kawasaki, J., S. Aegerter, R. D. Fevurly, A. Mammoto, T. Mammoto, M. Sahin, J. D. Mably, S. J. Fishman, and J. Chan. 2014. RASA1 functions in EPHB4 signaling pathway to suppress endothelial mTORC1 activity. *J. Clin. Invest.* 124:2774–2784.
- Keller-Costa, T., A. V. M. Canário, and P. C. Hubbard. 2015. Chemical communication in cichlids: A mini-review. *Gen. Comp. Endocrinol.* 221:64–74.
- Kettleborough, R. N. W., E. M. Busch-Nentwich, S. A. Harvey, C. M. Dooley, E. de Bruijn, F. van Eeden, I. Sealy, R. J. White, C. Herd, I. J. Nijman, et al. 2013. A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature* 496:494–497.
- Kocher, T. D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. *Nat. Rev. Genet.* 5:288–298.
- Koizumi, H., N. Yamaguchi, M. Hattori, T. O. Ishikawa, J. Aoki, M. M. Taketo, K. Inoue, and H. Arai. 2003. Targeted disruption of intracellular type I platelet activating factor-acetylhydrolase catalytic subunits causes severe impairment in spermatogenesis. *J. Biol. Chem.* 278:12489–12494.
- Lai, S. L., T. H. Chan, M. J. Huang, S. W. Low, and S. J. Lee. 2008. *PLoS One* 3:e3439.

- Lamichhaney, S., J. Berglund, M. S. Almén, K. Maqbool, M. Grabherr, A. Martinez-Barrio, M. Promerová, C.-J. Rubin, C. Wang, N. Zamani, et al. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518:371–375.
- Mabee, P., J. P. Balhoff, W. M. Dahdul, H. Lapp, P. E. Midford, T. J. Vision, and M. Westerfield. 2012. 500,000 fish phenotypes: The new informatics landscape for evolutionary and developmental biology of the vertebrate skeleton. *J. Appl. Ichthyol.* 28:300–305.
- Malinsky, M., R. J. Challis, A. M. Tyers, S. Schiffels, Y. Terai, B. P. Ngatunga, E. A. Miska, R. Durbin, M. J. Genner, and G. F. Turner. 2015. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science*. 350:1493–1498.
- Mallet, J., A. Meyer, P. Nosil, and J. L. Feder. 2009. Space, sympatry and speciation. *J. Evol. Biol.* 22:2332–41.
- Manda, P., J. P. Balhoff, H. Lapp, P. Mabee, and T. J. Vision. 2015. Using the phenoscape knowledgebase to relate genetic perturbations to phenotypic evolution. *Genesis* 53:561–571.
- Martin, C. H. 2012. Weak Disruptive Selection and Incomplete Phenotypic Divergence in Two Classic Examples of Sympatric Speciation: Cameroon Crater Lake Cichlids. *Am. Nat.* 180:E90–E109.
- Martin, C. H. 2013. Strong assortative mating by diet, color, size, and morphology but limited progress toward sympatric speciation in a classic example: Cameroon crater lake cichlids. *Evolution* 67:2114–23.
- Martin, C. H., J. S. Cutler, J. P. Friel, C. Denning Touokong, G. Coop, and P. C. Wainwright. 2015a. Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric speciation. *Evolution* 69:1406–1422.
- Martin, S. H., K. K. Dasmahapatra, N. J. Nadeau, C. Slazar, J. R. Walters, F. Simpson, M. Blaxter, A. Manica, J. Mallet, and C. D. Jiggins. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* 23:1817–1828.
- Martin, S. H., J. W. Davey, and C. D. Jiggins. 2015b. Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* 32:244–257.
- McGirr, J. A., and C. H. Martin. 2016. Novel candidate genes underlying extreme trophic specialization in Caribbean pupfishes. *Mol. Biol. Evol.* 34:873–888.
- McGirr, J. A., and C. H. Martin. 2018. Parallel evolution of gene expression between trophic specialists despite divergent genotypes and morphologies. *Evol. Lett.* 2-2: 62–75. <https://doi.org/10.1002/evl3.41>.
- McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernysky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, et al. 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297–1303.
- Meier, J. I., D. A. Marques, S. Mwaiko, C. E. Wagner, L. Excoffier, and O. Seehausen. 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* 8:14363.
- Midford, P. E., T. Dececchi, J. P. Balhoff, W. M. Dahdul, N. Ibrahim, H. Lapp, J. G. Lundberg, P. M. Mabee, P. C. Sereno, M. Westerfield, et al. 2013. The vertebrate taxonomy ontology: a framework for reasoning across model organism and species phenotypes. *J. Biomed. Semantics* 4:34.
- Muschick, M., M. Barluenga, W. Salzburger, and A. Meyer. 2011. Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. *BMC Evol. Biol.* 11:116.
- Musilová, Z., A. Indermaur, A. R. Bitja Nyom, R. Tropek, C. Martin, and U. K. Schlieven. 2014. Persistence of *Stomatepia mongo*, an endemic cichlid fish of the Barombi Mbo Crater Lake, Southwestern Cameroon, with notes on its life history and behavior. *Copeia* 14:556–560.
- Nachman, M. W., and B. A. Payseur. 2012. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos. Trans. R. Soc. B Biol. Sci.* 367:409–421.
- Nikaido, M., T. Ota, T. Hirata, H. Suzuki, Y. Satta, M. Aibara, S. I. Mzighani, C. Sturmbauer, K. Hagino-Yamagishi, and N. Okada. 2014. Multiple Episodic Evolution Events in V1R Receptor Genes of East-African Cichlids. *Genome Biol. Evol.* 6:1135–1144.
- Nikaido, M., H. Suzuki, A. Toyoda, A. Fujiyama, K. Hagino-Yamagishi, T. D. Kocher, K. Carleton, and N. Okada. 2013. Lineage-Specific Expansion of Vomeronasal Type 2 Receptor-Like (OlfC) Genes in Cichlids May Contribute to Diversification of Amino Acid Detection Systems. *Genome Biol. Evol.* 5:711–722.
- Noor, M. A. F., and S. M. Bennett. 2009. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* 103:439–444.
- Nosil, P., S. P. Egan, and D. J. Funk. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: “isolation by adaptation” and multiple roles for divergent selection. *Evolution* 62:316–36.
- Nosil, P., and J. L. Feder. 2012. Widespread yet heterogeneous genomic divergence. *Mol. Ecol.* 21:2829–2832.
- Papadopolus, S., A. S. T. Papadopoulos, W. J. Baker, D. Crayn, R. K. Butlin, R. G. Kynast, I. Hutton, and V. Savolainen. 2011. Speciation with gene flow on Lord Howe Island. *Proc. Natl. Acad. Sci. U. S. A.* 108:13188–93.
- Patterson, N., P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschorek, T. Webster, and D. Reich. 2012. Ancient admixture in human history. *Genetics* 192:1069–1093.
- Pickrell, J. K., and J. K. Pritchard. 2012. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genet.* 8:e1002967.
- Plenderleith, M., C. van Oosterhout, R. Robinson, and G. Turner. 2005. Female preference for conspecific males based on olfactory preference in a Lake Malawi cichlid fish. *Biol. Lett.* 1:1–8.
- Poelstra, J. W., N. Vijay, C. M. Bossu, H. Lantz, B. Ryll, I. Müller, V. Baglione, P. Unneberg, M. Wikelski, M. G. Grabherr, and J. B. W. Wolf. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* 344:1410–4.
- Poelstra, J.W., E.J. Richards, C.H. Martin. 2018. Speciation in sympatry with ongoing secondary gene flow and a potential olfactory trigger in a radiation of Cameroon cichlids. *Molecular Ecology*. Early view. <https://doi.org/10.1111/mec.14784>.
- Prescott, S. M., G. A. Zimmerman, D. M. Stafforini, and T. M. McIntyre. 2000. Platelet activating factor. *Annu. Rev. Biochem.* 69:419–445.
- Pritchard, J. K., J. K. Pickrell, and G. Coop. 2010. The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Curr. Biol.* 20:R208–R215.
- Ravinet, M., R. Faria, R. K. Butlin, J. Galindo, N. Bierne, M. Rafajlović, M. A. F. Noor, B. Mehlig, and A. M. Westram. 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J. Evol. Biol.* 30:1450–1477.
- Reich, D., K. Thangaraj, N. Patterson, A. L. Price, and L. Singh. 2009. Reconstructing Indian population history. *Nature* 461:489–94.
- Richards, E. J., and C. H. Martin. 2017. Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. *PLoS Genet.* 13:e1006919.
- Richards, E. J., M. R. Servedio, and C. H. Martin. 2018. New criteria for sympatric speciation in the genomic era. *bioRxiv*. <https://doi.org/10.1101/367623>.
- Schiffels, S., and R. Durbin. 2014. Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* 46:919.
- Schlieven, U. K., and B. Klee. 2004. Reticulate sympatric speciation in Cameroonian crater lake cichlids. *Front. Zool.* 1:5.
- Schlieven, U. K., D. Tautz, and S. Paabo. 1994. Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* 368:629–631.

- Schrider, D. R., F. K. Mendes, M. W. Hahn, and A. D. Kern. 2015. Soft Shoulders Ahead: Spurious Signatures of Soft and Partial Selective Sweeps Result from Linked Hard Sweeps. *Genetics* 200:267–284.
- Schumer, M., R. Cui, D. L. Powell, R. Dresner, G. G. Rosenthal, and P. Andolfatto. 2014. High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species. *Elife* 2014:1–21.
- Schumer, M., R. Cui, G. G. Rosenthal, and P. Andolfatto. 2015. Reproductive Isolation of Hybrid Populations Driven by Genetic Incompatibilities. *PLoS Genet.* 11:1–21.
- Schumer, M., C. Xu, D. L. Powell, A. Durvasula, L. Skov, C. Holland, J. C. Blazier, S. Sankararaman, P. Andolfatto, G. G. Rosenthal, and M. Przeworski. 2018. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*. <https://doi.org/10.1126/science.aar3684>
- Seehausen, O. 2013. Conditions when hybridization might predispose populations for adaptive radiation. *J. Evol. Biol.* 26:279–281.
- . 2004. Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19:198–207.
- Smukowski, C. S., and M. A. F. Noor. 2011. Recombination rate variation in closely related species. *Heredity* 107:496–508.
- Soria-Carrasco, V., Z. Gompert, A. A. Comeault, T. E. Farkas, T. L. Parchman, J. S. Johnston, C. A. Buerkle, J. L. Feder, J. Bast, T. Schwander, S. P. Egan, B. J. Crespi, and P. Nosil. 2014. Stick insect genomes reveal natural selection's role in parallel speciation. *Science* 344:738–42.
- Stankowski, S., and M. A. Streisfeld. 2015. Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proc. R. Soc. London B* 282:20151666.
- Stelkens, R. B., K. A. Young, and O. Seehausen. 2010. The accumulation of reproductive incompatibilities in African cichlid fish. *Evolution* 64: 617–33.
- The Heliconius Genome Consortium, K. K. Dasmahapatra, J. R. Walters, A. D. Briscoe, J. W. Davey, A. Whibley, N. J. Nadeau, A. V. Zimin, D. S. T. Hughes, L. C. Ferguson. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487:94–98.
- Trewavas, E., J. Green, and S. A. Corbet. 1972. Ecological studies on crater lakes in West Cameroon fishes of Barombi mbo. *J. Zool.* 167:41–95.
- Wagner, C. E., L. J. Harmon, and O. Seehausen. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* 487:366–369.
- Wu, C. I., and C. T. Ting. 2004. Genes and speciation. *Nat. Rev. Genet.* 5:114–122.
- Wu, C. I. 2001. The genic view of the process of speciation. *J. Evol. Biol.* 14:851–865.
- Yan, W., A. H. Assadi, A. Wynshaw-Boris, G. Eichele, M. M. Matzuk, and G. D. Clark. 2003. Previously uncharacterized roles of platelet-activating factor acetylhydrolase 1b complex in mouse spermatogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 100:7189–94.
- Zamani, N., P. Russell, H. Lantz, M. P. Hoepfner, J. R. Meadows, N. Vijay, E. Mauceli, F. Di Palma, K. Lindblad-Toh, P. Jern, et al. 2013. Unsupervised genome-wide recognition of local relationship patterns. *BMC Genomics* 14:347.

Associate Editor: Z. Gompert

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Percentages of the genome assigned to topologies under various filtering criteria.

Table S2. Overlap of *Saguaro* and *f_d* introgressed regions.

Table S3. Average *D_{xy}* among Barombi Mbo and riverine *Sarotherodon* species across the genome.

Table S4. Candidate introgressed regions in Barombi Mbo cichlid radiation that are shared across multiple species in a subclade.

Table S5. Candidate introgressed regions in Barombi Mbo cichlid radiation that are shared across multiple species across Barombi Mbo.

Table S6. Within-population genetic diversity in introgressed regions in Barombi Mbo and *Sarotherodon* riverine populations.

Figure S1. The percentage of the genome assigned to each topology by *Saguaro*.

Figure S2. *Saguaro* topologies featuring Barombi Mbo polyphyly with riverine populations involving the *Stomatepia* three-species complex.

Figure S3. *Saguaro* topologies featuring Barombi Mbo polyphyly with riverine populations involving the *Konia* + *Pungu* subclade.

Figure S4. *Saguaro* topologies featuring Barombi Mbo polyphyly with riverine populations involving the *Myaka* + *Sarotherodon* subclade.

Figure S5. The log-likelihood of *TreeMix* population graphs for Barombi Mbo cichlids as a function of the number of migration events.

Figure S6. Visualization of introgression with *S. galilaeus* MM across unassigned scaffolds for Barombi Mbo.

Figure S7. Visualization of introgression with *S. galilaeus* CR across linkage groups for Barombi Mbo.

Figure S8. Distribution of introgression block sizes detected from sliding window *f_d* statistic tests for Barombi Mbo species.

Figure S9. Visualization of divergence across the genome for species of *Stomatepia* and *S. galilaeus* populations from Mungo/Meme River (MM) and Cross River (CR).

Figure S10. Visualization of divergence across the genome for species of *Konia* and *S. galilaeus* populations from Mungo/Meme River and Cross River.

Figure S11. Visualization of divergence across the genome between *M. myaka* and *S. linnelli*, *S. galilaeus* populations from Mungo/Meme River, and Cross River.

Figure S12. Linkage disequilibrium decay among individuals used in this study.

Figure S13. Ancestral Population Size of Barombi Mbo Species.

Figure S14. Ancestral population size of outgroups lineages.